Vektron 6913 Fleet Trial Experimental Design and Analysis
Executive Summary

1.0 Experimental Design

A two-stage crossover design was selected. This design was selected instead of the standard parallel design because of its high efficiency and ability to reduce between vehicle (type) variations, and allow estimation of 'small' carryover effects. A priori carryover effects were expected to be small. In addition the expectation was that eight thousand miles would be a long washout period. The design can be found in the Vektron 1200 Gasoline Additive NOx Evaluation Fleet Test Program. The following describe the fuelling schemes and randomisation procedures.

1.1 Fueling Schemes

Scheme 1: SOTB (Start of Test Base for 1000 miles),
Run1 Base (Base for 8000 miles)
Run2 Addit (Additive in for 8000 miles either with constant additive (CA) or alternating fuel (AF))
Scheme 2: SOTB
Run1:Addit
Run2 Base

Scheme1= BBA, Scheme2= BAB

1.2 Random Vehicle Assignment

Vehicles were assigned to one of four fuelling strategies:

1. Scheme 1, CA
2. Scheme 2, CA
3. Scheme 1, AF
4. Scheme 2, AF

as vehicles became available by drawing from a hat. (See Vektron 1200 Gasoline Additive NOx Evaluation Fleet Test Program for the 7 vehicle types).

CA= Constant Additive, AF= Alternating Fuel

1.3 Random Driver Assignment

Two drivers were randomly assigned to drive half of the experimental runs. Post analysis found no significant driver affect.

1.4 Design Conclusions

It can be seen from the table above that there was a significant amount of effort taken to ensure that vehicle running assignment; driver, fuelling strategy, and running date were random. In fact the difficulty in finding vehicles (some from different states) added to the randomness of the process.

2.0 Data and Transformations

There is a significant amount of literature in the emissions area using either percent difference from SOTB, or natural log (ln) difference from the SOTB for those cases of only analysing Run1. Because this approach leads to some within vehicle correlation, Mixed-Effects models are designed to handle this kind of within group correlation. The decision was to do analysis using both percentage and natural log. These transformations gave normal residuals.

3.0 Outliers

There were only three outliers in this study. Two were mechanical and one statistical.

3.1 Mechanical Outliers

Two vehicles were found to have mechanical problems and were dropped. The vehicle coded FE-3 (Ford Escort) was caught early and was replaced with a comparable Escort coded as FE-5. In addition a vehicle coded as FF-4 had very high emissions (4 standard deviations away from the others) and was found to have a clogged EGR hose. This vehicle was dropped since it was already well into the test.

3.2 Statistical Outliers

When Run 1 data was modeled another vehicle was found to be a statistical outlier. This vehicle had very high oil consumption. This vehicle was also removed.

4.0 Analysis and Modeling

4.1 Data

Analysis was done on the Run 1 data prior to completion of Run 2. The reason was that these data represented a stand-alone parallel study. These analyses were done on the Wtd ftp from the FTP 75 test.

4.2 Test for covariates.

Miles: The plan was to remove miles effect when the data was modeled. However, the identification of a carryover effect made this impossible since there would be sever alliasing between affects. Looking only at the difference between the start of test base and Run1 removed this concern.

Initial Base ftp: In the case of NOx, the initial base weighted ftp were checked as a covariate and found to be very weekly correlated with the percent difference and log difference (R-sqr=0.33). This was not a significant covariate.

No statistical covariates could be found.

4.3 Constant Additive (CA) vs Alternating (AF) Fuelling

We tested if there was a difference between CA and AF. There was no difference. This was surprising since our A-priori expectation was that the AF fuel case would be somewhere between the Base and CA fuelling. We decided to combine the CA and AF into a variable called Additive Present. We felt this was justified and probably *conservative* since if AF were not the same as CA, it would have the effect of reducing the estimated effect.

4.4 Assignment of Terms

Using only Run1 and SOTB data, the model was very simple.

4.4.1 Random Effects

Key to the use of Mixed-Effects models is the assignment of the random effect group. To have balanced design and estimate variances correctly, each group member needs to receive **EVERY** treatment. In cases were an individual subject receives **EVERY** treatment, the group would be the individual. It is not possible in our case. The group therefore becomes vehicle type. The vehicle type (group) has all treatments (additive, miles, and order). Some missing cells are allowed (for example by dropping outliners) but need to be a small percentage of the data. There are many examples of this approach in the literature.

Infineum

4.4.2 Fixed Effects

Additive Present was the fixed effect term.

4.5 Tests For Carryover

As stated crossover designs are extremely efficient as long as there is no large carryover effect. There is a good deal of debate in the literature about what to do when there is a large carryover effect. Some say that no statistical analysis of the full (Run1 and Run2) dataset possible. The reason for this is there is no estimate of baseline in the second run. In our case carryover was severely alliased with miles. Others suggest a two-stage process: 1) if carryover is estimated and large, 2) the Run 1 data is used to determine the effect. When Run 2 was complete, tests were done for carryover. Because there was a significant carryover effect the decision was made to only use only the Run 1 data since it was an unbiased estimate of the effect and could be viewed as a stand alone parallel tests.

4.6 Analysis of Variance

4.6.1 NOx

The full analysis of variance and model output can be found in Appendix (A). The S-Plus procedure for Mixed Effects (lme) was used to model the data. The percent difference and log difference were so highly correlated (the correlation between the two prediction models was $R=.99$). The additive effect on NOx was found to be significant at ($p<0.06$) for both the log difference and percent difference. The estimated effect from the model for percent difference was 10.42 % (see model output).

Measure	**NOx**
p-value	**P<.06**

4.6.2 CO, HC, Fuel Economy (FE)

The full analysis of variance summary for these measures can also be found in Appendix (A). Only the log difference was tested for these measures. None of the above measures was found to be significant in Run 1. The table below summarizes the findings.

Measure	CO	HC	FE
p—value	P<0.39	P<0.33	P<0.21

Fuel economy was found to be directionally better.

5.0 References

1 Putt, M.., Chinchilli, V.M. (1999) "A Mixed Effects Model for the Analysis of Repeated Measure Cross-Over Studies." *Statistics in Medicine* 18 (22): Nov 30: 3037-3058

2 Freeman, P.R. (1989) "The Performance of the Two-Stage Analysis of Two-Treatment Two-Period Crossover Trials.", *Statistics in Medicine* 8 (12): 1421-1432

3 Woods, J. R., Williams, J.G., Tavel, M.. (1989) "The Two-Period Crossover Design in Medical Research." *Annals of Internal Medicine* 110 (7) : 560-565

4 Lindhjem, C.E. (EPA) (1995) "The Effect of Gasoline Reformulation and Sulfur Reduction on Exhaust Emissions from Post-1983 but Pre-1990 Vehicles" Gasoline Additives Emissions and Performance SAE SP-1095, Feb :97-142

5 Guerrieri, D.A., Caffrey, P.J., and Rao, V. (EPA) (1995) "Investigation into the Vehicle Exhaust Emissions of High Percentage Ethanol Blends" Gasoline Additives Emissions and Performance SAE SP-1095, Feb :85-103.

6 Korotney, D.J., Rao, V., Lindhjiem, C.E., and Sklar, M. S. (1995) "Reformulated Gasoline Effects on Exhaust Emissions: Phase III; Investigation of the effects of Sulfur, Olefins, Volatility, and Aromatics and Interactions Between Olefins and Volatility or Sulfur" Gasoline Additives Emissions and Performance SAE SP-1095, Feb: 179-186

7 Bitting, W.H, Firmstone, G.P., and Keller, C.,T. (1995) "A Fleet Test of Two Additive Technologies Comparing Their Effects on Tailpilpe Emissions, Biting, " Gasoline Additives Emissions and Performance SAE SP-1095, Feb: 61-84

8 Fleet Test Team "EPA-OMS, Original Equipment Makers and Infineum Discussions", *Internal Presentation given at EPA Ann Arbor*

9 Lehmacher, W. (1991) "Analysis of the Crossover Design in the Presence of Residual Effects.", *Statistics in Medicine* 10 (6). : 891-899.

10 Cotton, J. W. (1989) "Interpreting Data from Two-Period Crossover Design also Termed the Replicated 2x2 Lastin Square Design" Psychological *Bulletin* 106 (3).: 503-515.

11 Grieve, A.P. (1987) "A Note on the Analysis of the Two-Period Crossover Design when the Period-Treatment Interaction is Significant." *Biomedical Journal* 29 (7).: 771-775

12 Willian, A.R., Pater J.L. (1986) "Carryover and the Two-Period Crossover Clinical Trial." *Biometrics* 42 (33). : 593-600

13 Senn,S (1993) Cross-over Trials in Clinical Research, John Wiley & Sons, New York

14 Willian, A.R, Pater, J. L. (1986) "Using Baseline Measurements in the Two-Period Crossover Clinical Trial." *Controlled Clin Trials* 7(4).: 282-289

15 Hecker, H. (1986) "Identification and Interpretation of Effects in Two-Period Crossover Designs." *Edv in Medizin und Biologie* 17 (3). (RECD. 1987): 60-66

16 Grizzle, J. E. (1965) "Two-Period Change-Over Design and Its Use in Clinical Trials," *Biometrics*, v. 21: 467

17 Pineiro, J.C., Bates, D.M. (2000) Mixed-Effects Models in S and S-Plus, Springer-Verlag, New York

18 Vonesh,E.F., Chinchilli, V.M. (2000) Linear and Nonlinear Models for the Analysis of Repeated Measurement, Volume 154, Marcel Dekker, New York

Infineum

APPENDIX (A)

Model 1: Log difference in NOx.

Response: NODiff1= ln(Run 1 Wtdftp NOx)- ln(SOTB Wtdftp NOx)

```
Linear mixed-effects model fit by REML
 Data: Run1Diff.Matrix
        AIC       BIC   logLik
  -14.95505 -10.24283 11.47752

Random effects:
 Formula:  ~ 1 | VehTy
        (Intercept)  Residual
StdDev:   0.1017296 0.1135307

Fixed effects: NODiff1 ~ Run1A
                 Value  Std.Error DF   t-value p-value
(Intercept)  0.1439115 0.04985105 18  2.886830  0.0098
      Run1A -0.0908530 0.04493939 18 -2.021679  0.0583
 Correlation:
      (Intr)
Run1A -0.451

Standardized Within-Group Residuals:
       Min         Q1       Med        Q3      Max
 -1.547318 -0.7628003 0.1161634 0.6455034 2.002596

Number of Observations: 26
Number of Groups: 7
```

Model 2: Percent difference in NOx.

Response: PercNO= ((Run 1 Wtdftp NOx)- (SOTB Wtdftp NOx))/(SOTB Wtdftp NOx)

```
Linear mixed-effects model fit by REML
 Data: EPATestSet
        AIC       BIC   logLik
  -9.283391 -4.571175 8.641695

Random effects:
 Formula:  ~ 1 | VehType
        (Intercept)  Residual
StdDev:   0.1150019 0.1276391

Fixed effects: PercNO ~ Run1Additive
                  Value  Std.Error DF   t-value
 (Intercept)  0.1683925 0.05623054 18  2.994681
Run1Additive -0.1041735 0.05052549 18 -2.061802
```

```
            p-value
 (Intercept)  0.0078
Run1Additive  0.0540
 Correlation:
             (Intr)
Run1Additive -0.449

Standardized Within-Group Residuals:
       Min         Q1        Med        Q3      Max
 -1.423365 -0.7721178 0.07528225 0.6648652 2.088109

Number of Observations: 26
Number of Groups: 7
```

Model 3: Percent difference in CO.

Response: CODiff1= ln(Run 1 Wtdfp CO)- ln(SOTB Wtdftp CO)

```
Linear mixed-effects model fit by REML
 Data: Run1Diff.Matrix
        AIC       BIC   logLik
  -10.11854 -5.406327 9.059271

Random effects:
 Formula:  ~ 1 | VehTy
        (Intercept)  Residual
StdDev:  0.05631677 0.1406821

Fixed effects: CODiff1 ~ Run1A
                  Value  Std.Error DF    t-value
(Intercept) -0.02359255 0.04456802 18 -0.5293605
      Run1A  0.04937125 0.05541119 18  0.8909979
            p-value
(Intercept)  0.6030
      Run1A  0.3847
 Correlation:
      (Intr)
Run1A -0.622

Standardized Within-Group Residuals:
       Min        Q1         Med        Q3      Max
 -1.566994 -0.517042 -0.1081004 0.385755 2.786376

Number of Observations: 26
Number of Groups: 7
```

Model 4: Percent difference in HC.

Response: HCDiff1= ln(Run 1 Wtdfp HC)- ln(SOTB Wtdftp HC)

```
Linear mixed-effects model fit by REML
 Data: Run1Diff.Matrix
        AIC       BIC   logLik
  -35.69048 -30.97826 21.84524

Random effects:
 Formula:  ~ 1 | VehTy
        (Intercept)   Residual
StdDev:  0.02138314 0.08525533
```

```
Fixed effects: HCDiff1 ~ Run1A
                  Value  Std.Error DF   t-value
(Intercept) -0.02677977 0.02502756 18 -1.070011
      Run1A  0.03421637 0.03350816 18  1.021135
            p-value
(Intercept)  0.2988
      Run1A  0.3207
 Correlation:
      (Intr)
Run1A -0.669

Standardized Within-Group Residuals:
       Min         Q1         Med        Q3      Max
 -2.738632 -0.4886304 -0.01577301 0.6256249 1.959009

Number of Observations: 26
Number of Groups: 7
```

Model 5: Percent difference in Fuel Economy (FE).

Response: FEDiff1= ln(Run 1 Wtdfp FE)- ln(SOTB Wtdftp FE)

```
Linear mixed-effects model fit by REML
 Data: Run1Diff.Matrix
        AIC       BIC   logLik
  -122.8943 -118.1821 65.44715

Random effects:
 Formula:  ~ 1 | VehTy
        (Intercept)   Residual
StdDev: 0.009254287 0.01237101

Fixed effects: FEDiff1 ~ Run1A
                 Value   Std.Error DF  t-value p-value
(Intercept) 0.01033697 0.004916054 18 2.102697  0.0498
      Run1A 0.00650760 0.004891773 18 1.330315  0.2000
 Correlation:
      (Intr)
Run1A -0.498

Standardized Within-Group Residuals:
      Min         Q1         Med        Q3    Max
 -1.71365 -0.4958549 -0.08018999 0.7192495 2.1461

Number of Observations: 26
Number of Groups: 7
```